

CHILD LABOR AND THE EDUCATION OF A SOCIETY

CLIVE BELL

Südasiens-Institut der Universität Heidelberg

HANS GERSBACH

Eidgenössische Technische Hochschule Zürich

This paper analyzes policies by means of which a whole society in an initial state of illiteracy and low productivity can raise itself into a condition of continuous growth. Using an overlapping generations model in which human capital is formed through child rearing and formal education, we show that an escape from a poverty trap, in which children work full time and no human capital accumulation takes place, is possible through compulsory education or programs of taxes and transfers. If school attendance is unenforceable, temporary inequality is unavoidable if the society is to escape in finite time, but long-run inequalities are avoidable provided sufficiently heavy, but temporary, taxes can be imposed on the better off. Programs that aim simply at high attendance rates in the present can be strongly nonoptimal.

Keywords: Dynamic Redistribution Policies, Poverty Trap, Human Capital Formation, Temporary Inequality, Attendance Rates

1. INTRODUCTION

Parental care and abilities exert a powerful influence upon a child's linguistic and intellectual development. The process of child-rearing thus involves an intergenerational transfer that creates a certain potential in the child, which, if fully realized through the instilling of existing social knowledge, can lead to the continuous growth of human capital construed in the broad sense and hence of individual productivity.¹ Without formal education, however, children will almost surely attain only a small part of this potential, modest though it might be if their parents happen to be unlettered and poor. If children can work to supplement the family's income and parents cannot borrow against their children's future income [Baland and Robinson (2000), Hazan and Berdugo (2002)], parents may decide to deny their children any schooling at all. This perpetuates their own condition of

We thank François Bourguignon, Ramona Bruhns, Shanta Devarajan, Noemi Hummel, Bernhard Pacht, Martin Ravallion, Lars Siemers, and Nicholas Stern, as well as participants at the Development Economics Conference (NEUDC) 2000 and seminars at Memphis, Regensburg, and the World Bank, for helpful comments on earlier drafts. We are also indebted to two referees of this journal for valuable and constructive suggestions. The responsibility for all errors is ours alone. Address correspondence to: Hans Gersbach, CER-ETH—Center of Economic Research at ETH Zurich and CEPR, ETH Zurich, Zürichbergstrasse 18, 8092 Zurich, Switzerland; e-mail: hgersbach@ethz.ch.

ignorance and low productivity into the next generation, whereupon the process continues. From a long-term perspective, the ensuing failure to build human capital must be counted as the main social cost of child labor.

The purpose of this paper is to analyze policies by means of which a whole society that is initially mired in a state of illiteracy and low productivity can raise itself into a condition of literacy and continuous growth.² Our main innovation is the analysis of sequences of tax-and-transfer programs—as opposed to compulsory education.³ In view of the patchy enforcement of compulsory education in many poor countries, these programs offer a potentially more practicable way of escaping from the poverty trap.

We consider a model economy populated by overlapping generations with a continuum of agents who possess some altruism and choose their children's level of formal education. The formation of human capital occurs through a combination of child-rearing and formal schooling. Adults cannot borrow against their offspring's future income. All have the same preferences and the first generation start with identical endowments. Hence, any ensuing economic inequality, short-run or long-run alike, is caused by redistributive policies. The production technology is of the AK type.

The model yields the following results. First, an escape from the poverty trap can always be accomplished through a program of compulsory education, though lump-sum redistributive taxes may be necessary if the educational “technology” is not convex. Such a policy is effectively first-best in the setting considered here, because the allocation of current full income between current consumption and investment in human capital through education is directly determined. This benchmark optimum is also attainable through equivalent nonlinear subsidies and taxes.

If any of these instruments is not available, the situation takes on a second-best character. The intervention we analyze can be described as income support for some sections of the population, financed by lump-sum taxes on the rest. An escape from the poverty trap may still be possible, albeit with an important drawback: temporary inequality in post-tax incomes and human capital is a necessary condition to escape from poverty in finite time, though long-term inequalities can be avoided provided the government enjoys enough freedom in choosing taxes on incomes. An inability to tax the well-to-do sufficiently heavily during the transition can result in, not only a delay in attaining full, universal education, but also persistent long-run inequality.

Programs that minimize the time needed to attain this educational goal do not maximize school attendance rates in any particular period. Rather, they reflect the connection between the schooling received by succeeding members of the lineages of children who are educated today and the expansion of taxable capacity and the avoidance of subsidies in the future. Indeed, maximization of the school attendance rate in a particular period can induce a slide back into the poverty trap.

There is now a substantial literature on the economics of child labor.⁴ The possibility of multiple equilibria in a one-period setting, with a high-wage equilibrium

in which the children do not work, and a low-wage one in which they do, was first established by Basu and Van (1998) and discussed by Swinnerton and Rogers (1999). Some theoretical contributions employ a two-period structure. In Baland and Robinson (2000), parents expect to receive support from their children in old age, and their decisions concerning education can fall prey to time inconsistency. Ranjan (1999) implicitly rules out this particular problem, and concentrates instead on the case where parents would send their children to school if they had access to credit but are unable to borrow against the increase in income that schooling would generate. Jafarey and Lahiri (2000) examine to what extent credit opportunities such as “food for education” or “investment in education quality” can lower the incidence of child labor and increase the effectiveness of educational investments.

The possible connection between child labor and fertility is only touched on here. Dessy (2000) considers an infinitely lived “dynasty” with preferences over the sequences of consumption levels and family sizes. A poverty trap is possible, with an unstable equilibrium separating a low-productivity, high-fertility steady state from a high-productivity, low-fertility one. A poverty trap can also occur in Hazan and Berdugo’s (2002) OLG (overlapping generations) framework, in which educated labor is employed in an advanced sector, but children can only find traditional employment. Sustained growth in output per head is eventually triggered, and then maintained, by the accumulation of human capital, first through mere natural increase and then through the education of more slowly growing numbers, there being some critical level of aggregate human capital at which the investment in upgrading the technology needed to launch the advanced sector becomes profitable. Intervention in the form of compulsory education, with lump-sum taxation of the children on reaching adulthood to compensate their parents in old age, hastens this process and is Pareto-improving. Strulik (2004) highlights the role of child mortality in such processes of development.

Our normative analysis complements recent papers that provide full-fledged positive analyses of policies to combat child labor. Doepke and Zilibotti (2005) develop a model in which adults compete with children in the labor market and so may support a ban on child labor, unless their own working children provide a large fraction of family income. In this setting, a ban can be triggered by skill-biased technological change. This induces parents to choose smaller families. The model replicates features of the history of the United Kingdom in the 19th century. Krueger and Tjornhom (2005) examine a heterogeneous-agent general equilibrium model with a human capital externality. They show how the welfare consequences for individual households of a transition to policies that restrict child labor or provide tax-financed free education depend crucially on whether households have financial assets, or earn high or low wages. Finally, Galor and Moav (2006) show that increasing returns to human capital and the complementarity between physical and human capital in production create incentives for capitalists to design a tax policy that would generate universal schooling, and thus a reduction in child labor.⁵

Turning to empirical work on child labor and schooling, Grootaert and Kanbur (1995) and Maitra and Ray (2002) are comparative, cross-country studies. Edmonds (2000) points to a close relationship between child labor and home production, which can affect the policy recommendations drawn from models of child labor with parental preferences or credit constraints. Emerson and Souza (2003) provide evidence concerning the nature and persistence of child labor. There is also a substantial empirical literature on how child labor is connected with low income and lack of human capital accumulation; see, e.g., Grootaert and Kanbur (1995), Patrinos and Psacharopoulos (1997), Psacharopoulos (1997), Rosenzweig (1990), and Rosenzweig and Evenson (1977).

The plan of the paper is as follows. Section 2 lays out the basic model, which draws on Uzawa (1965) and is of the OLG variety. It starts with the technology for producing human capital and output, analyzes the household's behavior, and then derives the system's dynamics. Section 3 addresses the policy problem. The first-best solution can be attained by means of compulsory schooling, the precise schedule of which is derived for the case where the education technology is convex. The difficulties of implementing it are briefly discussed. All of the second-best solutions involve finding a program of taxes and income subsidies that will minimize the time needed for all households to attain a level of human capital such that either all children attend school full time or sustainable growth is ensured, subject to certain constraints on how much inequality can be tolerated. In Section 4, we analyze in detail the second-best, minimum-time program to educate the whole society when there is no limit on the degree of temporary inequality. The political economy of its implementation is taken up in Section 5. The paper concludes with a discussion of some aspects of development policy in the light of our findings and identifies a number of open issues.

2. THE BASIC MODEL

Consider an OLG model in which individuals live for two periods, which will be labeled childhood and adulthood, respectively. Each generation is the same size and consists of a continuum of households represented by $[0, 1]$.⁶ A household is indexed by i or j , where $i, j \in [0, 1]$. In the basic model, all households are alike and we drop indices.

In the absence of premature mortality, and assuming assortative mating, each household, or "family", may be taken to comprise one adult and one child (henceforth, mother and daughter). Let the proportion of childhood devoted to education in period t ($t = 0, 1, \dots$) be denoted by $e_t \in [0, 1]$, the residual being allocated to work. An adult is endowed with $1 + q$ units of time, of which the proportion $q/(1+q)$ is devoted to child-rearing and the remainder to work. For simplicity, it is assumed that these proportions are fixed. Given that one unit of time is devoted to work, let an adult in period t supply λ_t efficiency units of labor, where $\lambda_t \in [1, \infty)$ is a natural measure of her human capital, and the condition $\lambda = 1$ for the society as a whole can be thought of as a state of economic backwardness.

2.1. The Technology

Human capital is assumed to be formed through a process in which child-rearing is combined with formal education in the following way. In the course of rearing her daughter, the adult gives the child a certain capacity to build human capital for adulthood, the size of which is assumed to be related to the adult's own human capital. The adult's gift will be unavailing to preserve the child from the state of $\lambda = 1$ as an adult, however, unless it is complemented by some formal education. Hence, let the child's endowment of efficiency units of labor on reaching adulthood at time $t + 1$ be given by

$$\lambda_{t+1} = h(e_t) \cdot (z\lambda_t) + 1, \quad (1)$$

where z can be thought of as a transmission factor and the educational "technology" $h(\cdot)$ is assumed to be a continuous, strictly increasing, and differentiable function on $[0, 1)$, with $h(0) = 0$. Equation (1) implies that the gift of rearing and formal education are both necessary if human capital is to exceed the basic level $\lambda = 1$ in the next generation.⁷

For any sequence of formal education $\{e_t\}_{t=0}^{\infty}$, the intergenerational growth rate of the adult's human capital in a given household in period t , g_t , is given by

$$1 + g_t = \lambda_{t+1}/\lambda_t = zh(e_t) + (1/\lambda_t). \quad (2)$$

The level of λ is momentarily stationary; i.e., $\lambda_{t+1} = \lambda_t$, for all pairs (λ_t, e_t) satisfying

$$[1 - zh(e_t)]\lambda_t = 1. \quad (3)$$

One such pair is $(\lambda_t = 1, e_t = 0)$, namely, backwardness. In order to describe the evolution of human capital, which will be needed later, suppose that e_t takes the value of unity for all $\lambda_t \geq \lambda^a$. It is then seen from (2) that if the system starts from some value of $\lambda \geq \lambda^a$, and if the "technology" for (re)producing human capital is sufficiently productive, in the sense that it satisfies the condition $zh(1) \geq 1$, then λ_t will tend asymptotically to steady growth at the rate $[zh(1) - 1]$. In the special case $zh(1) = 1$, λ_t grows without limit, but its growth rate tends asymptotically to zero. If $zh(1) < 1$, however, then unbounded growth is impossible.

Having established these pregnant possibilities, we turn to the technology for producing output, which takes the form of an aggregate consumption good. With our sights on growth, and in view of the fact that human capital in the above account is a produced factor of production, let there be a proportional relationship between output and inputs of labor measured in efficiency units. All output will then accrue to the household as income.

The child's contribution to the household's income is given as follows: Without any education, the child will supply at most one efficiency unit of labor, because of the complementarity between the gift received during child-rearing and formal education. Indeed, it is plausible that the child's efficiency will be somewhat lower, *ceteris paribus*, on grounds of age alone. To reflect these considerations, let the

child be able to supply $\gamma(1 - e_t)$ efficiency units of labor when the child works $1 - e_t$ units of time. It is plausible to assume that $\gamma \in (0, 1]$; i.e., a child working full time is at most as productive as an uneducated adult. The household therefore supplies a total of $[\lambda_t + (1 - e_t)\gamma]$ efficiency units of labor to the production of the aggregate good. Under the above assumption on the technology, the level of output produced by a household that has endowment λ_t and chooses e_t is

$$y_t = \alpha[\lambda_t + (1 - e_t)\gamma], \quad (4)$$

where $\alpha \in (0, \infty)$ is the (constant) productivity of an efficiency unit of labor. Thus, we have a so-called AK model, the form of which has a certain affinity to Uzawa's (1965) learn-or-do model. Recalling the conditions for λ to grow (asymptotically) at a steady rate, it is seen at once that, should they hold, output per family will grow in the same manner if $e_t = 1 \forall t \geq t'$ for some t' .

2.2. The Household's Behavior

Following Basu and Van (1998), it is assumed that all allocative decisions lie in the adult's hands. We rule out any bequests at death, so that the whole of current income, as given by (4), is consumed. The gift of the factor $z\lambda_t$ through rearing is one form of inter vivos transfer. According to (1), however, the second form, namely, sending the child to school at least part of the time ($e_t > 0$), is also necessary if the child is to enjoy $\lambda_{t+1} > 1$ as an adult. Because current consumption is maximized by choosing $e_t = 0$, it follows that the adult's altruism toward her child must be sufficiently strong if she does choose $e_t > 0$.⁸

For simplicity, let the child's consumption be a fixed fraction $\beta \in (0, 1]$ of the adult's consumption c_t . From (4), we then obtain the family's budget line in the space of (c_t, e_t) :

$$(1 + \beta)c_t + \alpha\gamma e_t = \alpha(\lambda_t + \gamma). \quad (5)$$

Note that we assume that it is impossible for an adult to borrow against the child's future income, which has been justified extensively in the literature (see, e.g., Baland and Robinson [2000]). It will also be useful to define

$$\bar{c}(\lambda_t) \equiv \frac{\alpha(\lambda_t + \gamma)}{1 + \beta} \quad (6)$$

and

$$\underline{c}(\lambda_t) \equiv \frac{\alpha\lambda_t}{1 + \beta}, \quad (7)$$

which correspond to the consumption levels of the adult choosing $e_t = 0$ and $e_t = 1$, respectively. Given the endowment λ_t , her feasible set in the space of (c_t, λ_{t+1}) is the set $S(\lambda_t)$, defined as

$$\begin{aligned} S(\lambda_t) = \{ & (c_t, \lambda_{t+1}) : (1 + \beta)c_t + \alpha\gamma h^{-1}[(\lambda_{t+1} - 1)/z\lambda_t] \leq \alpha(\lambda_t + \gamma), \\ & 0 \leq c_t \leq \bar{c}(\lambda_t), 1 \leq \lambda_{t+1} \leq (z\lambda_t) \cdot h(1) + 1 \}. \end{aligned} \quad (8)$$

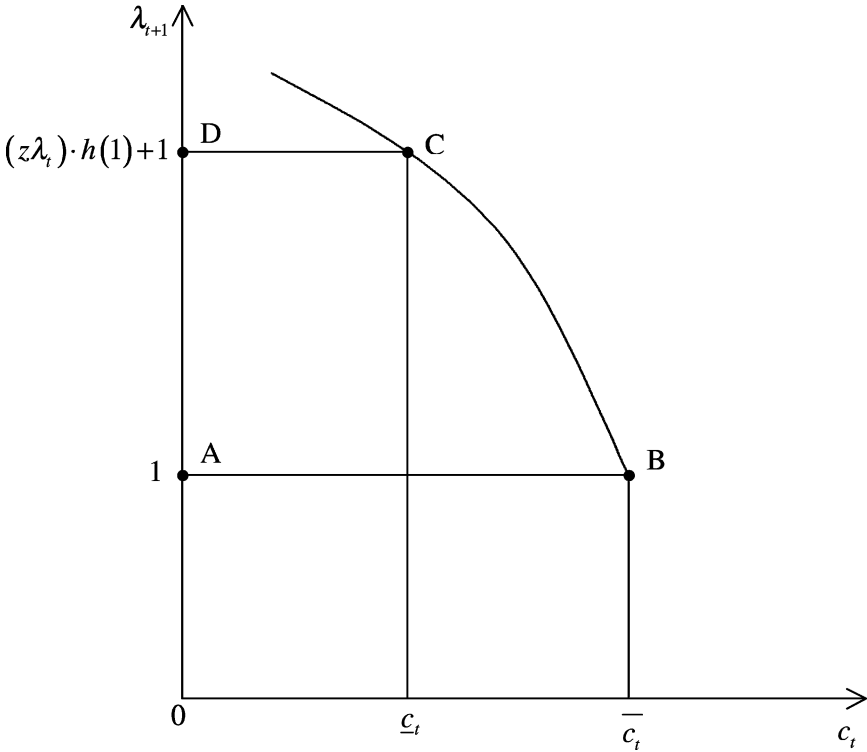


FIGURE 1. The household's feasible set.

The set $S(\lambda_t)$ is depicted in Figure 1 as $ABCD$, whose outer frontier BC is the locus generated by

$$\lambda_{t+1} = (z\lambda_t) \cdot h[(\alpha(\lambda_t + \gamma) - (1 + \beta)c_t)/\alpha\gamma] + 1, \quad c_t \in [\underline{c}(\lambda_t), \bar{c}(\lambda_t)] \quad (9)$$

and, purely for the purposes of illustration, $h(e_t)$ is assumed to be strictly concave.

It is clear from (6), (7), and (9) that an increase in λ_t enlarges the feasible set in such a way that it strictly dominates the set associated with the starting level of λ_t . There is also a substitution effect, for BC becomes steeper as λ_t increases. At any interior solution, it follows that λ_{t+1} is increasing in λ_t if λ_{t+1} is a normal good.

Let the adult's preference ordering be representable by the continuous, strictly increasing, differentiable, strictly quasi-concave function $u(c_t, \lambda_{t+1})$,⁹ and consider the problem

$$\begin{aligned} & \max_{(c_t, e_t)} \{u(c_t, \lambda_{t+1})\} & (10) \\ & \text{s.t. } [c_t, \lambda_{t+1}(e_t)] \in S(\lambda_t) \\ & e_t \in [0, 1] \end{aligned}$$

If $h(e_t)$ is a concave function, this problem will have a unique solution, denoted by $[c^0(\lambda_t), e^0(\lambda_t)]$, which is continuous in λ_t .¹⁰ In what follows, we often use the notation (c_t^0, e_t^0) as a convenient abbreviation for $[c^0(\lambda_t), e^0(\lambda_t)]$.

We make the following assumptions regarding the optimal choices (c_t^0, e_t^0) . First, if λ_t is very small, the family's existence will be so precarious that there may be no option but to put the child to work full time. At sufficiently high values, altruism will be operative, in the sense that the parent will choose $e_t > 0$. Therefore, we assume that there exists a critical value $\lambda^S > 1$ such that

$$\begin{aligned} [c^0(\lambda_t), e^0(\lambda_t)] &= [\bar{c}(\lambda_t), 0] \quad \forall \lambda_t \leq \lambda^S \\ e^0(\lambda_t) &> 0 \quad \forall \lambda_t > \lambda^S. \end{aligned} \quad (11)$$

The critical value λ^S is a limit that must be exceeded if altruism is to be operative.¹¹ Associated with λ^S is the level of consumption:

$$\bar{c}(\lambda^S) = \frac{\alpha(\lambda^S + \gamma)}{1 + \beta}. \quad (12)$$

Second, there exists a threshold value $\lambda^a > \lambda^S$ such that

$$[c^0(\lambda_t), e^0(\lambda_t)] = [\underline{c}(\lambda_t), 1] \quad \forall \lambda_t \geq \lambda^a. \quad (13)$$

The value λ^a can be thought of as marking the beginning of such affluence that the whole of childhood is spent at the school desk.

Third, it is assumed that the optimal choice e_t^0 is monotonically increasing in λ_t for all $\lambda_t \in (\lambda^S, \lambda^a)$.¹²

Before the difference equation (1) is analyzed, a remark should be made about the choice of the domain of the adult's preferences. An alternative would involve preferences over current consumption and the child's utility. In this case, the adult cares directly about the well-being of the child, though she knows that her daughter, as an adult, will care about her own child, and so forth. As in Barro (1974), therefore, all generations are effectively connected. In such a setting, by increasing the time devoted to schooling, the adult in period t enlarges her daughter's feasible set in such a way that the latter is always better off as a result, given that the daughter, too, will be confronted with the same problem of balancing current consumption in period $t + 1$ against her own daughter's well-being in period $t + 2$. Thus, each adult's utility depends upon the whole sequence of utilities achieved by her descendants. We expect that this alternative formulation would yield results qualitatively similar to those obtained here. Our formulation of preferences appears to be equally plausible and is much easier to handle in a technical sense.

2.3. Dynamics

Returning to (1) in the light of (11) and (13), we obtain

$$\lambda_{t+1}^i = \begin{cases} 1 & \forall \lambda_t^i \leq \lambda^S \\ zh[e^0(\lambda_t^i)]\lambda_t^i + 1 & \forall \lambda_t^i \in (\lambda^S, \lambda^a) \\ zh(1)\lambda_t^i + 1 & \forall \lambda_t^i \geq \lambda^a, \end{cases} \quad (14)$$

which yields the following results:

PROPOSITION 1.

- (i) $\lambda = 1$ is a locally stable equilibrium.
- (ii) If $zh(1) \geq 1$, then starting from any $\lambda \geq \lambda^a$, λ_t^i will grow without bound.
- (iii) If $zh(1)\lambda^a + 1 \geq \lambda^a$, there exists at least one stationary value of λ , denoted by λ^* , in $(\lambda^S, \lambda^a]$. Otherwise, there need be no such value.
- (iv) If $zh(1) < 1$ and $zh(1)\lambda^a + 1 > \lambda^a$, there exists a steady-state value of λ , $\lambda^{**} = 1/[1 - zh(1)] > \lambda^a$, such that λ^{**} is a stable, stationary equilibrium $\forall \lambda_t^i \geq \lambda^a$.

The proof of Proposition 1 is given in the Appendix.

Henceforth, the cases $zh(1) \geq 1$ and $zh(1) < 1$ will be called the *growth* and *nongrowth cases*, respectively.

The number of stationary values in the interval (λ^S, λ^a) , and hence the character of the system's dynamical behavior, depends on the shape of the function $h(e^0(\lambda_t))\lambda_t$. Although there is no space here to attempt a full characterization, a brief sketch will convey the flavor of the possibilities. Define

$$f(\lambda) \equiv h[e^0(\lambda)]\lambda, \quad \lambda \in (\lambda^S, \lambda^a),$$

where it is clear that $f(\cdot)$ is increasing in λ in the interval (λ^S, λ^a) . Differentiating twice, we have

$$f'' = 2h'e' + h''(e')^2\lambda + h'e''\lambda.$$

If $h'' \geq 0$ and $e'' \geq 0$, then f is convex everywhere on (λ^S, λ^a) . If, however, $h(\cdot)$ and $e^0(\cdot)$ are strictly concave, then $f'' < 0$ cannot be ruled out over the interval (λ^S, λ^a) . One possibility is that $f(\cdot)$ is first strictly concave and then strictly convex on (λ^S, λ^a) , so that there will be an odd number of stationary values in that interval in the growth case.

In order to see what all this portends for the behavior of λ_t , we illustrate the trajectory of λ_t for the growth case when there is a single stationary value of λ in the interval (λ^S, λ^a) (see Figure 2). The 45° line through the origin is labeled a ; the line g represents $\lambda_{t+1} = zh(1)\lambda_t + 1$, where g 's slope is $zh(1) \geq 1$. The horizontal segment BC reflects the fact that $\lambda_{t+1} = 1$ for all $\lambda_t \in [1, \lambda^S]$. The rising segment that passes through D and then meets g at λ^a arises from $\lambda_{t+1} = zh[e^0(\lambda_t)]\lambda_t + 1$ for all $\lambda_t \in (\lambda^S, \lambda^a)$. The instability of the stationary state corresponding to $\lambda_t = \lambda^*$ is evident from the trajectories drawn to the left and right of D, respectively.¹³

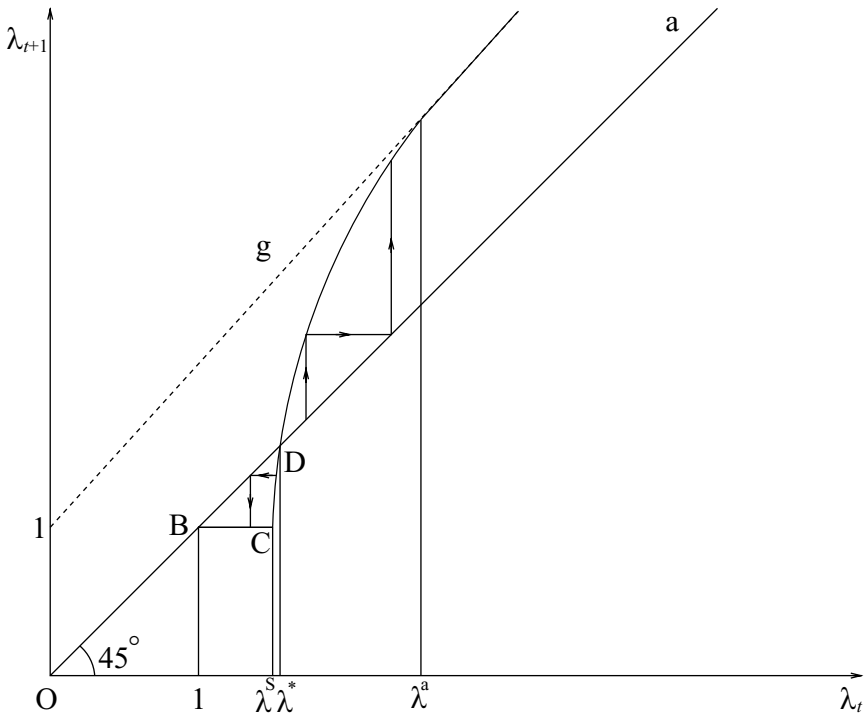


FIGURE 2. Growth case with two steady states.

3. THE POLICY PROBLEM

In this section, we formulate the policy problem associated with an initial state of backwardness, where the broad objective of policy is to liberate all lineages from this condition for good. The case for intervention in the present setting rests on the externalities that arise when the improvements in all future generations' welfare that would stem from better education of today's children are not fully reflected in the preferences of today's parents. This holds in our model because parents care about their children's human capital, but not about what happens subsequently. If, as is arguable, the government has a longer horizon than individual households, then the case for intervention to promote schooling at the expense of child labor is, in principle, established.¹⁴

The instruments available to the government for this purpose are assumed to be compulsory schooling and taxes and subsidies. The government is assumed to be able both to identify each household (or lineage), an ability that is vital to the payment of subsidies in an efficient way, and to assess its current level of full income. Because an adult's income is fixed in a particular period, a tax thereon is effectively lump-sum in nature, so that given the government's general objective, first-best allocations are, in principle, attainable. If compulsory schooling cannot

be enforced, or providing a program of sufficiently generous subsidies to attend school lies beyond the government's administrative and fiscal capacities, then one is driven to second-best solutions based on a combination of taxes and general income support. The effects of imposing further restrictions on the tax schedule will be briefly discussed after the main results have been derived.

In view of the emphasis on growth, we concentrate in what follows on the case where $zh(1) > 1$ and, for simplicity, there is a single stationary value of λ in the interval (λ^S, λ^a) , as depicted in Figure 2.

3.1. First-Best Policy

As a benchmark, consider the case where the government is able to determine directly how much each family consumes and the schooling its child receives, subject only to the requirement that each family obtain at least the subsistence minimum, which is defined to be c^{sub} for the consumption of adults and βc^{sub} for children. This minimum imposes an upper limit on the family's taxable capacity. We formulate the policy objective as minimizing the number of generations (periods) needed for the economy to attain a condition of self-sustaining growth in which compliance with full-time schooling occurs voluntarily. Formally, we define the policy problem P0 as

$$\begin{aligned} \text{P0 : } & \min_{\{(c_t^i, e_t^i)\}_{t=0,1,\dots}} \{T\} \\ \text{s.t. } & c_t^i \geq c^{\text{sub}}, e_t^i \in [0, 1], \quad \text{and} \quad \lambda_T^i \geq \lambda^a \forall i, t. \end{aligned}$$

We assume that the government can impose schooling and redistribute consumption goods subject to the aggregate resource constraint and $c_t^i \geq c^{\text{sub}}$. For the problem to be interesting, a family must have some taxable capacity, denoted by τ^{ba} , when $\lambda^i = 1$:

$$\tau^{ba} \equiv \alpha(1 + \gamma) - (1 + \beta)c^{\text{sub}} > 0, \quad (15)$$

where it is plausible that τ^{ba} is small, because households with $\lambda_t = 1$ may already be close to the subsistence level.

PROPOSITION 2. *If $h(e_t)$ is strictly concave, then starting from a condition of backwardness, the policy problem (P0) is solved by a monotonic program in which all children in each generation receive the same schooling, that is, $e_t^i = e_t^j \forall i, j$ and $\forall t \in \{0, \dots, T\}$, with $e_{t+1}^i > e_t^i$ if $e_t^i < 1$.*

The proof of Proposition 2 is given in the Appendix.

It is natural to ask whether the above program can be implemented by means of a suitable combination of taxes and subsidies. In view of the externality generated by education, let the government subsidize attendance at school and levy lump-sum taxes to finance the resulting outlays. Because all families remain identical under

the program, the index i may be dropped, and the budget line of any household in period t may be written as

$$(1 + \beta)c_t + (\alpha\gamma - \sigma_t)e_t = \alpha(\lambda_t + \gamma) - \tau_t,$$

where σ_t is the subsidy paid on each unit of schooling and τ_t is a poll tax. By setting $\sigma_t = \alpha\gamma \forall t$, the government will induce the family to choose $e_t^0 = 1$, for the net opportunity cost of education is then zero. The program must, however, be financially feasible without violating $c_t \geq c^{\text{sub}}$. In period 0, therefore, the subsidy rate $\sigma = \alpha\gamma$ is offered, up to a maximum payment of $\tau_0 = \tau^{ba}$. In such a situation, the family will choose exactly the optimum bundle derived in Proposition 2. The process is then continued for as many periods as needed, with τ_t being updated as λ_t grows.¹⁵

We now briefly discuss the character of the optimal program when $h(\cdot)$ is convex. If $h(\cdot)$ is linear, it is clear that the proof of Proposition 2 will still go through. If $h(\cdot)$ is strictly convex, however, then considerations of inequality may become relevant if (P0) is not soluble within one generation. For suppose h' is close to zero until e nears unity and then rises sharply: that is, education pays off handsomely only after a substantial threshold has been reached. In this case, it may well be better to educate some children fully in period 0 and then tax them heavily as adults in order to finance the education in period 1 of the offspring of those whose education was neglected in period 0. As we will see in the sections that follow, such schemes of redistribution across lineages are an essential feature of second-best programs based on general income support.

We close this analysis of the first-best with a discussion of the practical difficulties of implementing it. Many developing countries lack the central institutions and bureaucratic morale to ensure the truthful reporting of school attendance or to call the parents to book when the truants are reported. Large-scale “truancy” among teachers is also a problem, as the evidence in the World Development Report on the provision of public services [World Bank (2003)] vividly demonstrates. When compulsion is actually needed, there are correspondingly strong incentives for collusion between parents and teachers to exaggerate school attendance. The problems of enforcing compulsory schooling may also appear in connection with school attendance subsidies. The claims put in by numerous parents have to be verified frequently, and the incentives for teachers to report the children’s attendance truthfully are weak; but the strong incentives for collusion remain.¹⁶ In view of these difficulties, the remainder of the paper will be concerned with settings in which the first-best solution is not attainable, so that the government must resort to blunter instruments to achieve the goal set out above.

3.2. Second-Best Taxes and Subsidies

Observe from (8) and Figure 1 that a lump-sum subsidy will induce a parallel shift of BC to the right. The normality of λ_{t+1} implies that, for a given level of λ_t , such

a shift will result in an increase in e_t^0 for all $\lambda_t \in (\lambda^S, \lambda^a)$ and, if the subsidy is large enough, also for all $\lambda_t \in [1, \lambda^S]$. Hence the obvious next-best alternative is to provide families with what can be termed income support. This may take the form not only of direct transfers, if such are possible, but also of projects such as local infrastructure or the introduction of improved varieties of certain crops, whose benefits are reaped by particular groups or communities. As will become clear, second-best interventions lead to two characteristic results: first, the problem cannot be solved in one generation; and second, a solution, if one exists, always involves some inequality, if only temporarily.

For simplicity, we assume that taxes are solely a function of the income or human capital of adults.¹⁷ Let $\tau_t^i = \tau_t(\lambda_t^i)$ denote the tax levied in period t on household i , where $\tau_t^i \geq 0 \forall i, t$. Some fraction of the population will be subsidized out of the ensuing revenues. We denote by $s_t^i = s_t^i(\lambda_t^i)$ the subsidy household i will receive in period t if the adult has λ_t^i , where it should be noted that the subsidy depends on the index i directly, as well as indirectly through λ_t^i . The net income of household i in period t is

$$w_t^i = \alpha \lambda_t^i + \alpha(1 - e_t^i)\gamma + s_t^i(\lambda_t^i) - \tau_t(\lambda_t^i) \equiv w_t^{ia} + \alpha(1 - e_t^i)\gamma, \quad (16)$$

where w_t^{ia} denotes the net disposable income accruing to the adult in question. The household's net tax burden is defined by

$$v_t^i(\lambda_t^i) \equiv \tau_t(\lambda_t^i) - s_t^i(\lambda_t^i).$$

Throughout the remainder of the paper, households are either taxed or subsidized, and therefore $\tau_t(\lambda_t^i)$ and $s_t^i(\lambda_t^i)$ are net transfers. The evolution of human capital accumulation then follows the same logic as in Section 2.2. Recalling that an increase in λ_t both enlarges the set $S(\lambda_t)$ and makes its outer frontier BC everywhere steeper, the adult's choice of e_t^i is written as a function of both λ_t^i and v_t^i :

$$e_t^{i0} = e^0(\lambda_t^i, v_t^i).$$

Note that taxation need not differ across households that have the same taxable income. Subsidization, however, can and must be made dependent on income *and* the particular type of household. Although, in the end, only the net tax $v_t^i(\lambda_t^i)$ matters for household i , the distinction between taxation and subsidization will be useful in illustrating the working of different policies.

The maximum tax that can be obtained from any household is the difference between its full income and $(1 + \beta)c^{\text{sub}}$. The tax burden of household i is therefore assumed to be constrained by

$$\alpha \lambda_t^i - \tau_t(\lambda_t^i) + \alpha \gamma \geq (1 + \beta)c^{\text{sub}} \quad \forall i. \quad (17)$$

There may, of course, be further restrictions on the governments' ability to tax; for example, that post-tax income be everywhere strictly increasing in pretax income.

We assume that in drawing up programs, the society has to rely fully on its own fiscal resources.¹⁸ To formulate the budget constraints of the society as a whole, we reinterpret the indexation of households as a real-valued function on $[0, 1]$ that assigns every household its human capital in a particular period. Then denoting total government revenues in period t by B_t , the budget constraint is given by

$$B_t = \int_0^1 \tau_t(\lambda_t^i) di \geq \int_0^1 s_t^i(\lambda_t^i) di. \quad (18)$$

3.3. Policy Programs

A program of taxation and subsidization is to be chosen to bring the society out of backwardness in a sustainable way. There are several ways to formulate such a policy. Starting in period 0, let T denote the number of periods needed to bring all adults to at least the efficiency level λ^a .¹⁹ Because an increase in λ_t yields a strictly preferred set $S(\lambda_t)$, we work with a simple measure of inequality, namely²⁰

$$\Delta_t = \max_{i,j} (w_t^{ia} - w_t^{ja}). \quad (19)$$

All the formulations of the policy problem considered here involve minimizing T subject to some upper bound on the degree of inequality a society is prepared to tolerate. They can therefore be regarded as “turnpike” programs. The first is

$$\begin{aligned} \text{P1 :} \quad & \min_{\{(\tau_t(\lambda_t^i), s_t^i(\lambda_t^i))\}_{t=0,1,\dots}} \{T\} \\ \text{s.t.} \quad & \lambda_t \leq \bar{\Delta}, (17), (18), \text{ and } \lambda_T^i \geq \lambda^a \quad \forall i, t, \end{aligned}$$

where $\bar{\Delta}$ is the said upper bound on Δ_t . A special case of the policy problem (P1) is that where $\bar{\Delta} = \infty$, which seeks the fastest path unconstrained by inequality to the state in which all adults have at least the efficiency level λ^a . Observe that when all adults attain λ^a and none is subject to taxation, then all of their offspring will enjoy full-time schooling, and so attain at least λ^a .

Whereas policy problem (P1) focuses on inequality during the course of the program, others can be formulated that focus on the degree of long-run inequality that prevails after the program has been completed. Such an alternative policy problem can be stated as follows:

$$\begin{aligned} \text{P2 :} \quad & \min_{\{(\tau_t(\lambda_t^i), s_t^i(\lambda_t^i))\}_{t=0,1,\dots}} \{T\} \\ \text{s.t.} \quad & \Delta_T \leq \bar{\Delta}, (17), (18) \text{ and } \lambda_T^i \geq \lambda^a, \quad \forall i, t. \end{aligned}$$

Policy problem (P2) seeks the optimal speed program such that inequality is at most $\bar{\Delta}$ at the moment when the society becomes fully educated. An important special case of (P2) is $\bar{\Delta} = 0$, where no long-run inequality is allowed.

Other formulations are possible. In particular, we think that the following problem is relevant:

$$\begin{aligned} \text{P3 :} \quad & \min_{\{(\tau_t(\lambda_t^i), s_t^i(\lambda_t^i))\}_{t=0,1,\dots}} \{T\} \\ \text{s.t.} \quad & \Delta_t \leq \overline{\Delta}, (17), (18), \lambda_T^i \geq \lambda^a \text{ and } e_{t+1}^i \geq e_t^i, \forall i, t. \end{aligned}$$

In (P3) we impose the additional constraint on (P1) that the time spent in school by members of a particular lineage does not decline over time. On economic and on political economy grounds, this appears to be a sensible implicit restriction on the tax schedule. Imposing such a restriction can increase the time a society needs to escape backwardness.

3.4. The Inequality–Speed Dilemma

The above formulation of the policy problem involves a potential inequality–speed dilemma, which we now investigate. We assume that each household, viewed as a lineage, receives a positive subsidy only once.²¹ To simplify the exposition, we also assume that the tax levied on those households that are in a state of backwardness and are taxed is τ^{ba} .²² That being so, our first observation is obvious:

Fact 1. In an initial state of backwardness, equal treatment of citizens with respect to taxes and subsidies leaves the whole society in that state.

Equal treatment in period 0 would imply that each household would have a disposable income of $\alpha(1 + \gamma)$, which leads to $e_0^{i0} = 0$ and hence to $\lambda_1^i = 1 \forall i$. No household can escape from backwardness at any time under such a policy.

Having observed that creating inequality, if only temporarily, is a necessary condition to increase human capital, we now discuss how much inequality is needed if the whole society is to be educated. Because λ_t^i remains at unity until household i receives a subsidy and must reach at least λ^* in the period immediately afterward if it is not to revert to $\lambda^i = 1$, it follows that an escape from the poverty trap is possible only if there exists an s_t^i such that

$$zh[e^0(1, -s_t^i)] + 1 \geq \lambda^*. \quad (20)$$

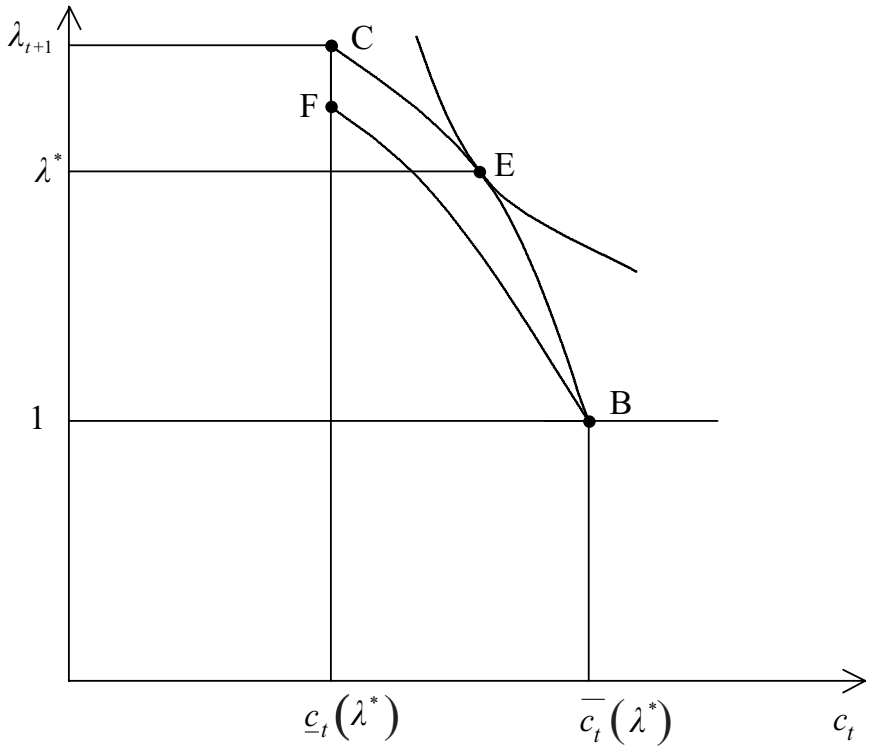
Let the smallest value of s_t^i that satisfies (20) be denoted by \tilde{s} , where it should be noted that \tilde{s} is independent of both i and t .

In period $t = 0$, the government's total revenue is $\tau^{ba}(1 - \delta_0)$, where δ_0 is the share of households that receive subsidies. The share of households that can be given the subsidy \tilde{s} is therefore

$$\tilde{\delta}_0 \equiv \frac{\tau^{ba}}{\tilde{s} + \tau^{ba}}. \quad (21)$$

It follows that

$$w_0^{ia} = \begin{cases} \alpha + \tilde{s} & i \in [0, \tilde{\delta}_0] \\ \alpha - \tau^{ba} & i \in (\tilde{\delta}_0, 1]. \end{cases}$$

FIGURE 3. The optimum when $\lambda = \lambda^*$.

Hence, when the subsidy is \tilde{s} , $\Delta_0 = \tilde{s} + \tau^{ba}$, which yields the following result:

PROPOSITION 3. *If $\bar{\Delta} < \tilde{s} + \tau^{ba}$ and $\lambda_0^i = 1 \forall i$, then the policy problem (P1) yields $T = \infty$.*

We now examine whether this result holds in the case where $\bar{\Delta} \geq \tilde{s} + \tau^{ba}$. In view of the above argument, the first step is to establish how large \tilde{s} is in relation to λ^* , where

$$\lambda^* = zh[e^0(\lambda^*, 0)]\lambda^* + 1. \quad (22)$$

This is depicted in Figure 3 as the point E on the frontier BC corresponding to $\lambda = \lambda^*$. Because $\bar{c}(\lambda^*) = \alpha(\lambda^* + \gamma)/(1 + \beta)$, it follows that the effect of giving a net transfer in the amount of $\alpha(\lambda^* - 1)$ to a household whose adult human capital is unity is to shift the outer frontier of its feasible set in such a way that point B is just feasible. The remainder of the said frontier, BF , lies below BC by virtue of $\lambda^* > 1$, which at once yields the result that the net transfer $\alpha(\lambda^* - 1)$ is too small

to induce $\lambda_{t+1}^i = \lambda^*$ when $\lambda_t^i = 1$. Hence,

$$\bar{s} > \alpha(\lambda^* - 1).$$

The special case where $\bar{\Delta} = \bar{s} + \tau^{ba}$ turns out to be of considerable importance; for it permits λ^* to be attained in the next generation, but no more than that level.

PROPOSITION 4. *If $\bar{\Delta} = \bar{s} + \tau^{ba}$ and $\lambda_0^i = 1 \forall i$, then the policy problem (P1) yields $T = \infty$.*

The proof of Proposition 4 is given in the Appendix.

Propositions 3 and 4 have profound implications for the tradeoff between speed and income inequality. The society must be prepared to tolerate a greater degree of inequality in incomes than $\bar{\Delta} = \bar{s} + \tau^{ba}$; and at some stage, some households must possess human capital in excess of λ^* while others are still in the state $\lambda = 1$, if all are eventually to escape from backwardness in finite time. Propositions 3 and 4 are in the nature of “impossibility” results. They evidently remain valid if the government’s ability to raise taxes is restricted beyond the basic requirement that all lineages can at least survive, as expressed by condition (17).

To complete this section, we need to say something about the case where $\bar{\Delta} > \bar{s} + \tau^{ba}$. This easing of the restriction on inequality permits $s_0^i > \bar{s}$, and hence $\lambda_1^i > \lambda^*$ for some i . Household i can now pay some taxes in period 1, while still choosing e_1^0 such that $\lambda_2^i > \lambda^*$ and without violating $\Delta_1 \leq \bar{\Delta}$. The taxable capacity so generated in period 1 undermines the argument in the proof of Proposition 4, and so opens up the possibility of a “bootstrap” operation, in which those lineages that receive a subsidy early on later pay the taxes needed to draw up the rest behind them. This possibility will now be explored in detail.

4. NO LIMITS ON TEMPORARY INEQUALITY

To focus on essentials, we begin by examining the special case of policy problem (P1) when no upper limit on temporary inequality is imposed ($\bar{\Delta} = \infty$). That is, we seek a policy yielding the swiftest possible attainment of a fully educated society. We assume that

$$zh(1) + 1 \geq \lambda^a. \quad (23)$$

Condition (23) states that if an adult in a state of backwardness were to educate her child fully, the child would choose full-time education for her own child, provided she herself were not taxed as an adult. This condition simplifies the analysis, but is not essential to our argument. It ensures that (P1) could, in principle, be solved in two periods: if those households singled out in period 0 for promotion from backwardness receive a subsidy sufficiently large to induce them to choose $e_0 = 1$, then their offspring will attain at least λ^a , and so provide a new tax base in period 1. If this tax base is large enough to subsidize all other households to the point where the latter choose $e_1 = 1$ in period 1, then (P1) will have been solved in two periods.

The tax–subsidy problem is complex and no closed-form solution exists that covers all possible periods T for which (P1) has a solution. Therefore, we proceed by considering the optimal policy for a given number of periods within which the entire society can be educated, ascending from $T = 1$. We establish conditions and optimal policies under the assumption that a particular period, say \bar{T} , is the outcome under (P1). If the solution does not yield \bar{T} , we move to $\bar{T} + 1$ and look once more. Starting from $\bar{T} = 1$ and using this method of induction, we will be able to characterize sequentially the solution under (P1). In the next section, we explore the case where all lineages can escape from backwardness within two generations. The extension to the case where the process takes three or more generations is set out in a technical appendix to Bell and Gersbach (2001). The consequences of the restrictions on the tax function $\tau_t^i(\cdot)$ implied by (P3) are taken up in Section 4.3.

4.1. Minimum Time: Two Periods

We first note that $T = 1$ can never be the outcome under (P1), because in $t = 0$ only a fraction of the society can receive positive net transfers such that the adults in $t = 1$ will have achieved λ^a . The first possibility, then, is to escape from backwardness in two periods. In $t = 0$, suppose that the share of households with $i \in [0, \delta_0]$ will be subsidized. The total subsidy to this group of families is given by

$$s_0^i = \frac{1 - \delta_0}{\delta_0} \tau^{ba} \equiv s_0, \quad (24)$$

which implies the following pattern of human capital formation:

$$\lambda_1^i = \begin{cases} zh[e^0(1, -s_0)] + 1 & \text{if } i \in [0, \delta_0] \\ 1 & \text{if } i \in (\delta_0, 1]. \end{cases} \quad (25)$$

To solve the problem in two periods, the tax–subsidy scheme in $t = 1$ must fulfill

$$\lambda_2^i = \begin{cases} zh\{e^0[\lambda_1^i, \tau_1(\lambda_1^i)]\} \lambda_1^i + 1 \geq \lambda^a & \text{if } i \in [0, \delta_0] \\ zh[e^0(1, -s_1^i)] + 1 \geq \lambda^a & \text{if } i \in (\delta_0, 1] \end{cases} \quad (26)$$

and the financing constraint

$$\delta_0 \tau_1 \geq (1 - \delta_0) s_1, \quad (27)$$

where the superscript i may be dropped without ambiguity in (27) by virtue of the fact that both groups are homogeneous. For each value of λ_1^i ($i \in [0, \delta_0]$), let $\hat{\tau}_1(\lambda_1^i)$ denote the maximum tax that household i can pay without violating the upper branch of (26). Because λ_1^i is uniquely determined by δ_0 through s_0 , it follows that for each choice of $\delta_0 \in [0, 1]$, we obtain a unique λ_1^i ($i \in [0, \delta_0]$) and a unique $\hat{\tau}_1(\lambda_1^i)$. For each s_1 that satisfies the lower branch of (26), $\{\delta_0 \hat{\tau}_1[\lambda_1^i(\delta_0)] - (1 - \delta_0) s_1\}$

is also a continuous function. Consider, therefore, the following problem:

$$\max_{0 \leq \delta_0 \leq 1} (\delta_0 \hat{\tau}_1 \{\lambda_1^i[s_0(\delta_0)]\} - (1 - \delta_0)s_1) \quad i \in [0, \delta_0]. \quad (28)$$

Because the objective function is continuous and $\delta_0 \in [0, 1]$, a solution exists and is denoted by $\hat{\delta}_0$. We obtain

PROPOSITION 5. *If $\hat{\delta}_0 \hat{\tau}_1(\lambda_1^i(s_0(\hat{\delta}_0))) - (1 - \hat{\delta}_0)s_1 \geq 0$, then $T = 2$ is the outcome of policy problem (P1) with $\overline{\Delta} = \infty$. In particular, the subsidy $\hat{s}_0 = (1 - \hat{\delta}_0)\tau^{ba}/\hat{\delta}_0$ to the households $i \in [0, \hat{\delta}_0]$ in period zero ensures that $T = 2$.*

The preceding analysis highlights the point that two considerations enter into the determination of $\hat{\delta}_0$ or, equivalently, of \hat{s}_0 . First, how large is the taxable capacity in the next period yielded by subsidizing a fraction of households in the present period, a capacity expressed by $\delta_0 \tau_1 \{\lambda_1^i[s_0(\delta_0)]\}$? Second, how large is the burden of future subsidies needed to promote the rest, as expressed by $(1 - \delta_0)s_1$?

Proposition 5 carries a number of further implications. First, if $T = 2$ under (P1), there will not, in general, be a unique optimal tax and subsidy policy. For suppose that the budget surplus in the second period is positive, $\hat{\delta}_0 \tau_1 \{\lambda_1^i[s_0(\hat{\delta}_0)]\} - (1 - \hat{\delta}_0)s_1 > 0$, and denote by $\bar{\delta}_0$ one value of the share of subsidized persons in the first period for which the above financing constraint, (27), holds as an equality.²³ Then there exists a continuum of solutions for δ_0 and $s_0(\delta_0)$ with $\delta_0 \in [\bar{\delta}_0, \hat{\delta}_0]$ and s_0 given by (24).

Second, to resolve this nonuniqueness, additional considerations can be brought into the reckoning. An obvious choice is to minimize inequality. Let attention be confined, to start with, to the first period ($t = 0$). Because $\Delta_0 = s_0 + \tau^{ba}$, the problem is as follows:

$$\min_{\delta_0 \in [0, 1]} s_0 \quad \text{s.t. (24)–(27)}. \quad (29)$$

Observe from (25) that minimizing s_0 entails minimizing λ_1^i ($i \in [0, \delta_0]$) and hence also τ_1 . It is seen from (24) that minimizing s_0 is also equivalent to maximizing δ_0 . It follows that if problem (29) possesses a solution, it will be unique. Denote the solution by δ_0^0 , so that

$$s_0^0 = \frac{1 - \delta_0^0}{\delta_0^0} \tau^{ba}.$$

Now recall from the lower branch of (26) that when the subsidy s_1 is the minimum net transfer needed to yield $\lambda_2 = \lambda^a$ for the group supported in period 1, it imposes the smallest fiscal burden in period 1 if the number of families to be promoted is also a minimum, from which it also follows that the tax in period 1 on each family promoted in period 0, τ_1^i ($i \in [0, \delta_0]$), is a minimum, too. That the solution to (29) entails minimizing the number of families to be promoted in period 1 therefore motivates

PROPOSITION 6. *If the solution of policy problem (P1) with $\overline{\Delta} = \infty$ yields $T = 2$, then the solution of problem (29), δ_0^0 , yields minimum inequality in both periods, with no inequality ($\lambda_2^i = \lambda^a \forall i$) in the second.*

The proof of Proposition 6 is given in the Appendix.

It should be remarked that the argument will go through perfectly well even if the upper frontiers are not concave everywhere, provided the optimum bundle (c_1^0, λ_2^0) is always unique.

The third implication of Proposition 5 is that although income inequality is temporarily created in order to escape from the state of low productivity, there need be no inequality in human capital, and hence no inequality in future income, at the end of the two-period program. A subset of the programs in Proposition 5 and the program in Proposition 6 therefore have the desirable feature that although initially identical households are treated differently with respect to taxes and subsidies in the first period, their offspring become identical once more in the next period, and enter the stage of continuous growth as such. Hence, the apparently highly restrictive constraint $\Delta_T = 0$ is not binding for (P2) if $T = 2$ is the outcome of (P1). This observation can easily be generalized to the case where $T \geq 3$, as we now show.

4.2. An Equivalence between (P1) and (P2)

We claim that if (17) is the only restriction placed on taxes, then second-best, maximal-speed programs can always avoid long-run inequality, in the sense that all lineages attain the same level of human capital at the conclusion of such programs. The following corollary is proved in the Appendix.

COROLLARY 1. *Suppose that $\{\tau_t^*(\lambda_t^i), s_t^{*i}(\lambda_t^i)\}$ is a solution of (P1) for $\overline{\Delta} = \infty$, with minimum time T . Then T is also the minimum time of (P2) with $\Delta_T = 0$.*

4.3. Persistent Inequality

The preceding corollary shows that long-run inequalities can be avoided, but it implies that the schooling of children of educated adults in period $T - 1$ must be so limited that their human capital as adults will be exactly λ^a in period T . Yet this might be politically implausible or even impossible, because the implied range of income over which the marginal tax rate is 100% could be very large. Policy problem (P3) addresses this difficulty by imposing the constraint that e_t be nondecreasing within each lineage. This is equivalent to imposing certain restrictions on the tax function $\tau_t^i(\cdot)$. It will now be shown by means of an example that the optimal program for policy problem (P3) can require more time to attain the goal of λ^a for the whole population than problem (P2), and may result in income inequalities that persist beyond the program's horizon.

Consider the case of symmetric Stone–Geary preferences in the form

$$u = \begin{cases} (c - g)(z\lambda \cdot h(e) + 1) + (g - c^{\text{sub}}) & \text{if } c \geq g \\ c - c^{\text{sub}} & \text{if } g \geq c \geq c^{\text{sub}} \\ -\infty & \text{otherwise.} \end{cases}$$

Consumption must be at least c^{sub} for “survival,” and full income per equivalent adult must exceed the amount g if any schooling is to occur. Let the educational technology take the form $h(e) = b \cdot e$, and consider the following constellation of parameter values: $\alpha = 0.16$, $\beta = 1$, $\gamma = 0.7$, $c^{\text{sub}} = 0.1$, $g = 0.14$, $z = 0.7$ and $b = 5.05$. As established in Bell and Gersbach (2001), this example yields the claimed result.

4.4. An Algorithm for Maximal Speed

We complete this section by generalizing the procedure for finding programs that will bring about a fully educated society in the least possible time.

The previous sections indicate how maximal speed programs have to be designed. Ascending from $T = 1$, one can establish necessary and sufficient conditions that the least time be a particular period, say T . We summarize the procedure as follows:

CHARACTERIZATION OF MAXIMAL SPEED PROGRAMS

A least-time optimal program starting from $t = 0$ spans T periods if and only if there exists a sequence of tax-and-transfer schemes such that

- (i) all households attain at least λ^a in period T ;
- (ii) there is no budget deficit in period $T - 1$ and there are balanced budgets in all previous periods;
- (iii) no such program exists for $T - 1$ periods.

A formal description of the algorithm is given in Bell and Gersbach (2001). Although the principle underlying the procedure for deriving speed programs is outlined above, there are further considerations when we move to longer time-horizons.

An important feature of maximal speed programs is that school attendance rates in all but the last period are not maximized. Indeed, the preceding results imply that distributing subsidies more broadly to increase the number of children at school in any particular period can lead to a return to poverty in the future, because the succeeding cohorts of these lineages will not attend school to the same extent, if at all. A second feature of such programs is that the indeterminacy of tax-and-transfer schemes increases as \bar{T} increases, which opens the door to the use of additional criteria for judging inequality.

4.5. Endogenous Fertility

As demonstrated by Raut and Srinivasan (1994), endogenous fertility can significantly affect the dynamics of OLG models. Allowing households to choose fertility (n_t) in our framework, with preferences represented by a utility function $u(c_t, n_t, \lambda_{t+1})$, would introduce the familiar quantity–quality tradeoff. Our sequences of tax–subsidy schemes can be adapted to this case, but at the cost of a tremendously expanded technical apparatus. The main consequences of such an extension for our results can be summarized as follows.

Suppose fertility increases with the adult's income, so that the dependency ratio rises, too.²⁴ Then the required subsidy to promote the escape from the poverty trap becomes larger and, initially at least, fewer households can receive subsidies, which reduces the speed toward a fully educated society. However, the fact that such households have more children implies a larger future tax base, so it increases the speed with which the remaining poor households will be pulled up. The acceleration will continue over the following generations. Exactly the opposite effects occur when fertility declines with rising income. Thus, introducing fertility decisions does not affect the basic logic of our dynamic tax/subsidy schemes, but it does introduce further static and dynamic feedback effects that impinge on the speed with which optimal programs achieve full education.

5. POLITICAL IMPLEMENTATION

In this section, we briefly address the positive question of whether second-best policies are likely to be implemented. Whether overcoming poverty can and will occur under democracy will depend crucially on the specific constitutional rules for setting the agenda and taking decisions. When, as in this paper, the level of an adult's human capital is fixed by past decisions, a democracy with simple majority rule and equal agenda-setting rights will not achieve universal literacy and growth. To establish this claim, suppose a subset of households start to educate their children. If they form a minority, these families will be taxed at very high rates in the next voting round, when uneducated individuals will form a majority. As a consequence, although an initial education-enhancing redistribution may occur (the winners might be chosen by lottery), its effect will be destroyed through time. If, conversely, the said subset form a majority, then they will be able to block any redistributive measures in favor of the minority left in illiteracy and poverty in the first round.

There are, however, democratic constitutions with specific sets of rules that do enable a democracy to produce an educated society. A constitution involving the simple-majority rule, a tax-protection rule, and a rotating agenda-setting rule may achieve this objective.²⁵ The tax-protection rule, if it limits marginal tax rates sufficiently, ensures that educated households do not fall back into poverty. Rotating agenda-setting ensures that all uneducated individuals will have the power to determine the agenda at some point in time. As agenda setters can channel

subsidies to themselves, such a rule guarantees that all individuals eventually will be educated. Of course, the speed of human capital accumulation will be lower under a democracy than under a benevolent dictatorship, as a majority must always be better off in every voting round, which constrains the set of education-enhancing proposals.

It should be noted that the tax-protection rule may arise as the outcome of a cooperative game. If adults can effectively destroy some or all of their endowments, by emigrating or not working, for example, then there are limits on how punitive taxation can be. This threat is an essential element in Aumann and Kurz's (1977) analysis of taxation under simple majority voting. One of their salient results is that net income is an increasing function of gross income, with the poor receiving some income support. Another is that the marginal tax rate will be at least 50%, but will never reach 100%. Such an endogenous limitation has something in common with our policy program (P3). Although all members of society will eventually enjoy a full education, initial inequalities will not necessarily disappear, even in the long run.

6. CONCLUSIONS

This investigation of the nexus of child labor, economic growth, and inequality rests largely on the assumption that parents decide on their children's schooling, an assumption that can be justified by the enormous difficulties of monitoring and enforcing school attendance and labor standards [see, for example, Basu and Van (1998) and Basu (1999)]. Our analysis has yielded several findings, while leaving a number of issues unsettled.

The main conclusions are, first, that a whole society or group can be mired in a stable state of ignorance and low productivity, wherein all children work full-time to supplement their families' income and, in missing out on schooling, thereby perpetuate as adults the condition into which they themselves were born. Second, an escape from this poverty trap into a state of universal literacy and continuous economic growth is possible through various second-best programs of redistributive taxes and income support, even without outside aid. Third, the catch here is that in the absence of outside aid on a sufficiently large scale, such programs necessarily involve some temporary inequality, and they may result in long-run inequality if the government is unable to tax the better-off sufficiently heavily. For the very limited taxable capacity of the economy in its original state demands that the subsidies thus financed, which are designed to induce voluntary schooling, must be chosen in such a way as to create additional taxable capacity in the future. This implies that the beneficiaries' offspring must attain a level of education such that their lineages do not slide back into the poverty trap. Fourth, the necessity of temporary inequality speaks against policies aimed at bringing about high attendance rates in the short term without giving due consideration to the long-term perspective within which the problem is to be solved.

The possibility of compulsory schooling as a complement to taxes and subsidies raises some interesting issues where the optimal package of measures is concerned. Suppose, for example, that child labor can be limited to the point where the family can just afford the subsistence level of consumption in the absence of any taxation of income. If the educational technology is convex, a sequence of partial bans will yield an escape from the trap in minimum time without resort to taxes and subsidies. If the educational technology is not convex, however, it is not clear that, when the goal is to bring about a fully educated society in minimum time, universal compulsory schooling would eliminate the necessity of some degree of temporary inequality. For inequality can increase taxable capacity, and taxes and subsidies are also needed to attain the optimum in this case. These matters will be important in situations where some monitoring and enforcement are possible under governmental or international standards.

The list of open issues does not end with those arising from compulsory education. First, what other concepts of inequality, both within and across generations, commend themselves in the evaluation of alternative paths to full literacy? Second, how should foreign aid be allocated in such a setting? Third, how do the necessity of employing distortionary taxes and the possibility of tax evasion affect the complexion, or even the feasibility, of an optimal program? Fourth, how can credit opportunities for individual households or for the country as a whole help to decrease the time needed to educate a society? These and other pressing issues in the area of child labor and education are left as topics for further research.

NOTES

1. Becker et al. (1990) and Ehrlich and Lui (1991) pioneered this approach based on the direct transmission of potential productivity from parent to child.

2. The paper is therefore related to some recent literature on economic growth in which the transition through different regimes is endogenously generated. The initial contributions are those of Galor and Weil (1999, 2000), who describe, within a unified framework, long-run development processes from an epoch of Malthusian stagnation to a state of sustained economic growth in modern times.

3. There is a large literature on growth and inequality, surveyed in Aghion and Williamson (1998). This does not, however, deal with the problems of finding a path from one steady state to another, as we do here. The work of Galor and Zeira (1993) shows that the distribution of wealth matters for economic growth. At this fundamental level, our paper deals with a similar question to that in the growth and inequality literature: do short-term or long-run tradeoffs exist between growth and inequality?

4. For an early survey, see Basu (1999).

5. In our model, output is produced by means of human capital alone, so there are no capitalists.

6. We comment on the robustness of our results to endogenous fertility in Section 4.5.

7. The evolution of human capital described in equation (1) is the simplest formulation, because it does not involve any persistence. Persistence can be introduced by extending equation (1) to $\lambda_{t+1} = h(e_t)(z\lambda_t) + 1 + \rho\lambda_t$. For ρ sufficiently small, our analysis should still hold. In future research, we hope to say more about the case where persistence is large.

8. An alternative possibility is that children are expected to support their parents in old age, with the contributions becoming more generous with the rising level of the children's income as adults. To pursue this possibility, three overlapping generations would be needed.

9. Note that the utility function is defined on the set $\{(c_t, \lambda_{t+1}) : c_t \geq 0, \lambda_{t+1} \geq 1\}$.

10. If, however, $h(\cdot)$ is strictly convex on some interval—for example, when e is small and the payoff to additional schooling has yet to become large—the solution need not be unique and the locus $(c^0(\lambda_t), e^0(\lambda_t))$ may exhibit a discontinuity at one or more values of λ_t . This possibility is a distraction from the main theme of the paper and is ruled out by assumption, whereby it should be remarked that it is not necessary that $h(\cdot)$ be concave for the solution to be unique.

11. This is equivalent to assuming that, for all $\lambda_t \leq \lambda^s$, the indifference curves are sufficiently steep where they intersect the horizontal line through $\lambda_{t+1} = 1$.

12. The assumption holds if λ_{t+1} is a sufficiently strong luxury good for all $\lambda_t > \lambda^s$, and if $h(e_t)$ is not too strongly concave. A detailed discussion and justification of the assumption is contained in our working paper [Bell and Gersbach (2001)].

13. There are further possibilities associated with the non-growth case, which are discussed in Bell and Gersbach (2001).

14. A detailed discussion of the case for intervention is contained in Bell and Gersbach (2001).

15. Given that the authorities can both observe and condition taxes on e_t , it is also possible to devise a nonlinear income tax schedule that will induce families to choose $e_t^0 = \min[\xi_t/2, 1]$ without any taxes actually being paid.

16. Some decentralized schemes for poor communities, such as Progresa in Mexico, appear to have overcome the worst of these problems, though at considerable cost to the central fisc. Whether they can be replicated in other, poorer countries is open to question.

17. This may be justified by the ease of tax evasion for child income. It is unlikely that allowing household income to be taxable would change the main results that follow.

18. The case where a fixed amount of outside aid is available is discussed in Bell and Gersbach (2001).

19. As will become clear, the policy problems are formulated in such a way that this level is also indefinitely sustainable after completion of the programs in question. The requirement to bring the whole society to $(\lambda^* + \epsilon)$ for some $\epsilon > 0$ would yield the same qualitative result.

20. Inequality will arise only from differences across groups, so that taking the range of the distribution as a measure of inequality is defensible. The qualitative nature of our results appears to be robust with respect to more sophisticated measures of inequality.

21. The proofs of the propositions rely crucially on this assumption. Although the overall conclusions are robust if we drop it, the proofs become extremely tedious, and the boundaries for the inequality measure need to be slightly altered.

22. This assumption permits one to avoid considering the case where poor adults are taxed by an arbitrarily small amount. Without it, the formulation of Propositions 3 and 4 becomes a bit more cumbersome.

23. There may be other solutions. In that case, we take one for which the budget surplus is positive in the interval $[\bar{\delta}_0, \hat{\delta}_0]$.

24. It is straightforward to give examples in which fertility can increase or decrease when income rises from very low levels.

25. Such rules are discussed in Gersbach and Siemers (2005).

26. Equivalent is the case where the population is divided into two and each half is treated symmetrically.

27. This claim follows from the first-order conditions for the said expression to attain a maximum, whereby $\psi_1^1 = \psi_1^2$, which will hold only by choosing equal treatment in period 0.

REFERENCES

- Aghion, P. and J.G. Williamson (1998) *Growth, Inequality and Globalization: Theory, History and Policy*. New York and Cambridge, UK: Cambridge University Press.
- Aumann, R.J. and M. Kurz (1977) Power and taxes. *Econometrica* 45(5), 1137–1161.
- Baland, J.-M. and J. Robinson (2000) Is child labor inefficient? *Journal of Political Economy* 108(4), 663–679.

- Barro, R. (1974) Are government bonds net wealth? *Journal of Political Economy* 82, 1095–1117.
- Basu, K. (1999) Child labor: Cause, consequence, and cure, with remarks on international labor standards. *Journal of Economic Literature* 37, 1083–1119.
- Basu, K. and P.H. Van (1998) The economics of child labor. *American Economic Review* 88, 412–427.
- Becker, G.S., K.M. Murphy, and R. Tamura (1990) Human capital, fertility, and economic growth. *Journal of Political Economy* 98, S12–S37.
- Bell, C. and H. Gersbach (2001) Child Labor and the Education of a Society. IZA Working Paper No. 338.
- Dessy, S.E. (2000) A defense of compulsive measures against child labor. *Journal of Development Economics* 62, 261–275.
- Doepke, M. and F. Zilibotti (2005) The macroeconomics of child labor regulation. *American Economic Review* 95(5), 1492–1524.
- Edmonds, E. (2000) The Relationship Between Child Labor and Home Production: Evidence from Variation in Sibling Composition. Mimeo, Dartmouth College.
- Ehrlich, I. and F.T. Lui (1991) Intergenerational trade, longevity, and economic growth. *Journal of Political Economy* 99, 1029–1059.
- Emerson, P. and A.P. Souza (2003) Is there a child labor trap? Intergenerational persistence of child labor in Brazil. *Economic Development and Cultural Change* 51, 375–398.
- Galor, O. and O. Moav (2006) Das Human-kapital: A theory of the demise of the class structure. *Review of Economic Studies* 73(1), 85–117.
- Galor, O. and D. Weil (1999) From Malthusian stagnation to modern growth. *American Economic Review* 89, 150–154.
- Galor, O. and D. Weil (2000) Population, technology, and growth: From the Malthusian regime to the demographic transition and beyond. *American Economic Review* 90, 806–828.
- Galor, O. and J. Zeira (1993) Income distribution and macroeconomics. *Review of Economic Studies* 60, 35–52.
- Gersbach, H. and L.-H.R. Siemers (2005) Can Democracy Educate a Society? IZA Discussion Paper No. 1693.
- Grootaert, C. and R. Kanbur (1995) Child labour: An economic perspective. *International Labour Review* 134(2), 187–203.
- Hazan, M. and B. Berdugo (2002) Child labor, fertility, and economic growth. *Economic Journal* 112, 810–828.
- Jafarey, S. and S. Lahiri (2000) Food for Education and Funds for Education Quality: Policy Options to Reduce Child Labour. Mimeo, University of Wales, Swansea, and University of Essex.
- Krueger, D. and J. Tjornhom (2005) On the distributional consequences of child labor legislation. *International Economic Review* 46(3), 785–815.
- Maitra, P. and R. Ray (2002) The joint estimation of child participation in schooling and employment: Comparative evidence from three continents. *Oxford Development Studies* 30(1), 41–62.
- Patrinós, H.A. and G. Psacharopoulos (1997) Family size, schooling and child labor in Peru: An empirical analysis. *Journal of Population Economics* 10, 387–405.
- Psacharopoulos, G. (1997) Child labor versus educational attainment: Some evidence from Latin America. *Journal of Population Economics* 10, 377–386.
- Ranjan, P. (1999) An economic analysis of child labor. *Economic Letters* 64, 99–105.
- Raut, L.K. and T.N. Srinivasan (1994) Dynamics of endogenous growth. *Economic Theory* 4, 777–790.
- Rosenzweig, M. (1990) Population growth and human capital investments: Theory and evidence. *Journal of Political Economy* 98, S38–S70.
- Rosenzweig, M. and R. Evenson (1977) Fertility schooling and economic contribution of children in rural India: An econometric analysis. *Econometrica* 45, 1065–1079.
- Strulik, H. (2004) Child mortality, child labour, and economic development. *Economic Journal* 114, 547–568.
- Swinerton, K.A. and C.A. Rogers (1999) The economics of child labor: Comment. *American Economic Review* 89(5), 1382–1385.

Uzawa, H. (1965) Optimum technical change in an aggregative model of economic growth. *International Economic Review* 6, 18–31.

World Bank (2003) *World Development Report 2004: Making Services Work for Poor People*. New York: Oxford University Press.

APPENDIX

PROOF OF PROPOSITION 1

Part (i) follows immediately from the assumption that $\lambda^S > 1$. Part (ii) follows at once from the third part of (14). To prove part (iii), observe that because $zh(e^0(\lambda_t^i))\lambda_t^i$ is a continuous function, it follows from $\lambda_{t+1}^i = 1 \forall \lambda_t^i \leq \lambda^S$ and $\lambda_{t+1}^i(\lambda^a) = zh(1)\lambda^a + 1 \geq \lambda^a$ that λ_t possesses at least one stationary value in the interval $(\lambda^S, \lambda^a]$. If, however, $zh(1)\lambda^a + 1 < \lambda^a$, there need be no such value. In the case where $zh(1)\lambda^a + 1 > \lambda^a$ and $zh(1) < 1$, the largest stationary value of λ_t is $1/[1 - zh(1)] > \lambda^a$, which is clearly stable $\forall \lambda_t^i \geq \lambda^a$. ■

PROOF OF PROPOSITION 2

Given the nature of the claim, it clearly suffices to consider just two households ($i = 1, 2$), whose level of human capital may differ after period 0.²⁶ Define $\Lambda_t \equiv \lambda_t^1 + \lambda_t^2$, namely, their aggregate human capital at time t . Because their aggregate consumption must be at least $2(1 + \beta)c^{\text{sub}}$, the aggregate time their children spend at school must satisfy

$$e_t^1 + e_t^2 \leq [\Lambda_t + 2\gamma - 2(1 + \beta)c^{\text{sub}}/\alpha]/\gamma \equiv \xi_t.$$

In period 0, $\lambda_0^1 = \lambda_0^2 = 1$. If $\xi_0 = 2[(1 + \gamma) - (1 + \beta)c^{\text{sub}}/\alpha]/\gamma \geq 2$ and $zh(1) + 1 \geq \lambda^a$, then by choosing $e_0^1 = e_0^2 = 1$, the problem will have been solved in one generation. If, however, $zh(1) + 1 < \lambda^a$, both Λ_1 and ξ_1 will have been maximized in this case, with $\lambda_1^1 = \lambda_1^2 = zh(1) + 1$ in period 1.

The other, more plausible, alternative in the state of backwardness is $\xi_0 < 2$, so that universal, full-time, compulsory schooling is not feasible in period 0. Note that the assumption $\alpha(1 + \gamma) > (1 + \beta)c^{\text{sup}}$ ensures that $\xi_0 > 0$. In this case,

$$\Lambda_1 = z[h(e_0^1) + h(e_0^2)] + 2 = z[h(e_0^1) + h(\xi_0 - e_0^1)] + 2,$$

which is maximized by choosing $e_0^1 = e_0^2 = \xi_0/2$ by virtue of the strict concavity of $h(\cdot)$. If $zh(\xi_0/2) + 1 \geq \lambda^a$, the problem will have been solved in one generation. Otherwise, we proceed to period 1.

Suppose the policy of equal treatment is continued in period 1. If the policy program $e_t^1 = e_t^2$ ($t = 0, 1$) yields the maximal value of Λ_2 and the latter is at least $2\lambda^a$, then the problem will have been solved in two generations. Whatever the policy chosen in period 0, however, we have

$$\Lambda_2 = \begin{cases} z[h(e_1^1)\lambda_1^1 + h(\xi_1 - e_1^1)\lambda_1^2] + 2 & \text{if } \xi_1 < 2 \\ z[h(1)\lambda_1^1 + h(1)\lambda_1^2] + 2 & \text{if } \xi_1 \geq 2. \end{cases}$$

Define $\psi_t^i \equiv \lambda_t^i / \Lambda_t$. Then Λ_2 may be rewritten as the product of two functions:

$$\Lambda_2 = \begin{cases} z\Lambda_1 \left[h(e_1^1) \psi_1^1 + h(\xi_1 - e_1^1) \psi_1^2 \right] + 2 & \text{if } \xi_1 < 2 \\ z\Lambda_1 \left[h(1) \psi_1^1 + h(1) \psi_1^2 \right] + 2 & \text{if } \xi_1 \geq 2. \end{cases}$$

Consider the upper branch first. Because identical treatment of children in period 0 maximizes Λ_1 , to prove that Λ_2 is maximized by a policy of equal treatment in both periods, it suffices to show that the expression $[h(e_1^1) \psi_1^1 + h(\xi_1 - e_1^1) \psi_1^2]$ is maximized by such a policy. Because identical treatment in period 0 also maximizes ξ_1 , it follows from the strict concavity of $h(\cdot)$ that the expression $[h(e_1^1) \psi_1^1 + h(\xi_1 - e_1^1) \psi_1^2]$ is maximized by choosing identical treatment in period 1 and $\psi_1^1 = \psi_1^2$.²⁷ Following inspection of the lower branch, we conclude that identical treatment of children within each period maximizes Λ_2 . If, moreover, the said value of Λ_2 is at least $2\lambda^a$, then the problem will have been solved in two generations.

By writing Λ_t as

$$\Lambda_t = \begin{cases} z\Lambda_{t-1} \left[h(e_{t-1}^1) \psi_{t-1}^1 + h(\xi_{t-1} - e_{t-1}^1) \psi_{t-1}^2 \right] + 2 & \text{if } \xi_{t-1} < 2 \\ z\Lambda_{t-1} \left[h(1) \psi_{t-1}^1 + h(1) \psi_{t-1}^2 \right] + 2 & \text{if } \xi_{t-1} \geq 2, \end{cases}$$

it is clear that exactly the same argument will hold over as many periods as necessary to yield $\lambda_T^i \geq \lambda^a$ ($i = 1, 2$). ■

PROOF OF PROPOSITION 4

In period 0, $\bar{\Delta} = \tilde{s} + \tau^{ba}$ implies $s_0^i = \tilde{s}$ and $\tilde{\delta}_0 = \tau^{ba} / (\tilde{s} + \tau^{ba})$. Now consider the economy in $t = 1$. The fraction $\tilde{\delta}_0$ of the adult population has human capital λ^* . If they are to choose a level of education such that their children attain at least λ^* , then they cannot be taxed. It is clear from Figure 2 and (22) that any net transfer from these households to the state would yield $\lambda < \lambda^*$ in the next generation. Therefore, to move a fraction of the remaining households, denoted by $\tilde{\delta}_1$, to the knowledge level λ^* in the next period ($t = 2$), the government's budget constraint implies that

$$\tilde{\delta}_1(1 - \tilde{\delta}_0)\tilde{s} = (1 - \tilde{\delta}_0)(1 - \tilde{\delta}_1)\tau^{ba}, \quad (\text{A.1})$$

and from (21) we obtain $\tilde{\delta}_0 = \tilde{\delta}_1$. Repeating our argument for every period yields

$$\tilde{\delta}_t = \tilde{\delta}_{t-1} = \dots \tilde{\delta}_0 \equiv \tilde{\delta}. \quad (\text{A.2})$$

Therefore, after any finite number of periods, say n , a positive measure $(1 - \tilde{\delta})^n$ of households in the population remains in the state of backwardness, which in turn implies $T = \infty$. ■

PROOF OF PROPOSITION 6

Consider Figure 4, in which the upper frontier of the feasible set of a family promoted by a subsidy s_1 in period 1 is drawn as $B'C'$, to which an indifference curve is tangential at C' . (In the absence of the subsidy, the said frontier is BC , of which $B'C'$ is a parallel shift.) A family promoted in period 0 attains λ_1^1 in period 1 and pays the tax τ_1^1 . Its frontier is drawn

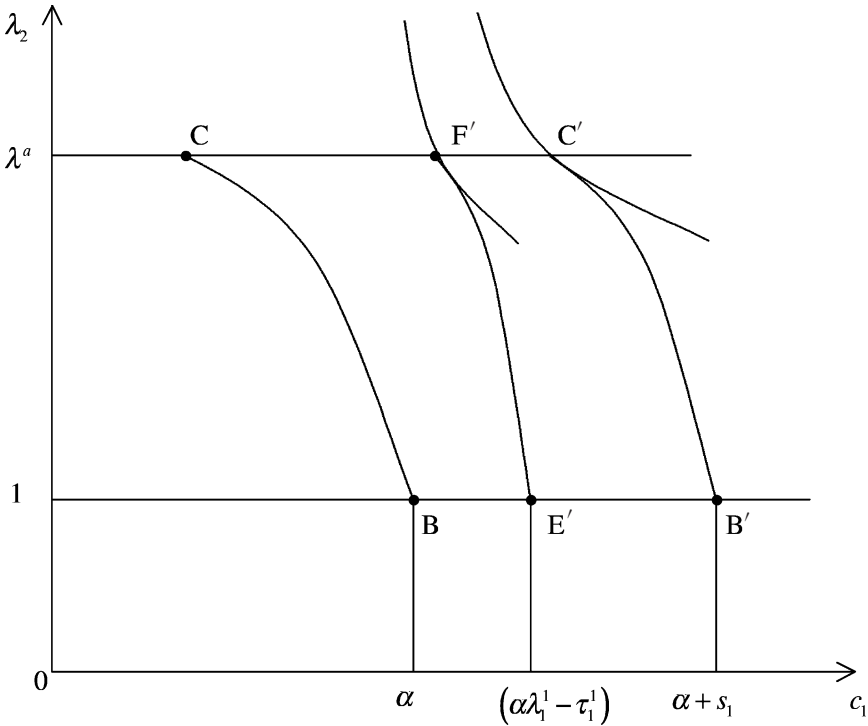


FIGURE 4. Minimum time for $T = 2$.

as $E'F'$; by virtue of the upper branch of (26), an indifference curve is tangential to it at F' . Because λ_2^i is a normal good, the indifference curves become flatter as one proceeds along the horizontal line $\lambda_2 = \lambda^a$. Hence, if the group promoted in period 0 is to attain exactly $\lambda_2 = \lambda^a$, it follows that $E'F'$ must lie to the left of $B'C'$, where it will be recalled that, for each λ_2 , $E'F'$ is steeper than $B'C'$ by virtue of $\lambda_1^1 > 1$. It then follows that

$$\begin{aligned}\Delta_1 &= |\alpha\{zh[e^0(1, -s_0)] + 1\} - \tau_1^1(s_0) - (\alpha + s_1)| \\ &= s_1 + \tau_1^1(s_0) - \alpha zh[e^0(1, -s_0)].\end{aligned}$$

Hence, to minimize Δ_1 , we must minimize $\{\tau_1(s_0) - \alpha zh[e^0(1, -s_0)]\}$. It is seen from Figure 4 that this is equivalent to minimizing $E'B'$, whereby B' is fixed by s_1 . Now, an increase in s_0 will increase λ_1^1 and hence make $E'F'$ steeper, thereby requiring that E' be pushed to the left in order to maintain $\lambda_2^1 = \lambda^a$. A reduction in s_0 is not feasible, for the financing constraint must be satisfied in period 1. ■

PROOF OF COROLLARY 1

Because T is feasible under (P1) with $\bar{\Delta} = \infty$, all households will have attained at least λ^a at the end of T periods. Suppose household i attains $\lambda_T^i > \lambda^a$ after T periods, having

reached λ_{T-1}^i in period $T - 1$. Because $zh(e_t)\lambda_t$ is continuous and monotonic in e_t and λ_t , and $zh(0)\lambda + 1 = 1 < \lambda^a$, there exists a tax $\bar{\tau}_{T-1}(\lambda_{T-1}^i) > \tau_{T-1}^*(\lambda_{T-1}^i)$ such that $\lambda_T^i = \lambda^a$. Because additional taxation of this kind increases the funds available to subsidize the poor without jeopardizing the condition $\lambda_T^i \geq \lambda^a$, the solution of (P1) with the modified tax scheme still yields T . Because these considerations apply to any household that attains more human capital than λ^a in period T , the tax scheme $\bar{\tau}_{T-1}(\lambda_{T-1}^i)$ can be made independent of a particular household i . Therefore, by appropriate taxation in period $T - 1$, all households will reach λ^a in period T , which proves the corollary. ■